Lecture notes – Integrators

Contents

1 Integrating Ordinary Differential Equations
   1.1 Euler’s method ................................................................. 1
   1.2 Using the midpoint and the Runge-Kutta method .................... 4
      1.2.1 The 4th order Runge-Kutta method ............................ 6
   1.3 Adaptive Step-Sizes ......................................................... 6
   1.4 More than one degree of freedom and dynamics with a potential force ... 8
   1.5 The meaning of incompressible flow ................................... 9
   1.6 Volume conservation in Phase Space ................................... 11
   1.7 Comparison between symplectic and non-symplectic first order integration for the Harmonic oscillator ................................................ 12
   1.8 Comparison between 2-nd order Runge Kutta and leap-frog midpoint methods 14
   1.9 Integrating many particles ................................................. 15
   1.10 Burlisch-Stoer Method ..................................................... 15
   1.11 Low Order Symplectic integrators – Störmer-Verlet (leapfrog) ........... 15

1 Integrating Ordinary Differential Equations

1.1 Euler’s method

We start with a dynamical system described by the equation

\[ \dot{x} = f(x,t) \]

Starting with initial condition \( x_0 \) at \( t_0 \) we would like to advance and compute an approximate solution for \( x \) at a later time \( t = t_0 + h \). The ideal solution would be a function \( x(t) \) but we only start known the slope of this function at \( x_0, t_0 \). We could expand this ideal solution using a Taylor series

\[
x(t_0 + h) = x(t_0) + x'(t_0)h + x''(t_0)\frac{h^2}{2} + x'''(t_0)\frac{h^3}{3!} + \ldots
\]  \hspace{1cm} (1)

Here the function \( x(t) \) with \( x(t_0) = x_0 \) and the derivatives are with respect to time. We note that the first order term has slope \( x'(t_0) \) that is equal to \( f(x_0, t_0) \). Ignoring second order terms

\[
x(t_0 + h) \sim x(t_0) + hf(x_0, t_0).
\]

This is known as the Euler method and it is first order which means that the error is second order or proportional to \( h^2 \).

How large is the error? Looking at the Taylor series and comparing equation to equation 1 the error has size

\[
\text{error} \sim x''(t_0)h^2/2.
\]

It makes sense that the error depends on the second derivative because if the actual solution was linear the Euler method would have given the correct answer.

This depends on the actual solution as a function of time.

\[
x''(t_0) = \frac{d}{dt} x'(t_0) = \frac{d}{dt} f(x, t) \bigg|_{t_0}
= \frac{\partial f}{\partial x} \frac{dx}{dt} \bigg|_{t_0} + \frac{\partial f}{\partial t} \bigg|_{x_0, t_0}
= \frac{\partial f}{\partial x} \bigg|_{x_0, t_0} f(x_0, t_0) + \frac{\partial f}{\partial t} \bigg|_{x_0, t_0}
\]

The error for the first order method depends on the slope of \( f \) and \( h^2 \). As \( f \) gives the time derivative, the second derivative of \( x(t) \) depends on the first derivative of \( f \).
Figure 1: Illustrating using the Euler method to integrate an ordinary differential equation.
1.2 Using the midpoint and the Runge-Kutta method

Let us consider expanding the solution for \( x(t) \) around time \( t = t_0 + h/2 \), the midpoint in time!

\[
\begin{align*}
    x(t_0 + h) &= x(t_0 + h/2) + \left. \frac{h}{2} \frac{dx}{dt} \right|_{t_0 + h/2} + \left. \frac{h^2}{8} \frac{d^2x}{dt^2} \right|_{t_0 + h/2} + O(h^3) \\
    x(t_0) &= x(t_0 + h/2) - \left. \frac{h}{2} \frac{dx}{dt} \right|_{t_0 + h/2} + \left. \frac{h^2}{8} \frac{d^2x}{dt^2} \right|_{t_0 + h/2} + O(h^3)
\end{align*}
\]

Subtracting these two equations gives

\[
x(t_0 + h) = x(t_0) + h \left. \frac{dx}{dt} \right|_{t_0 + h/2} + O(h^3)
\]

\[
= x(t_0) + hf(x(t_0 + h/2), t_0 + h/2) + O(h^3)
\]

Notice that the second order terms cancel. To use this we need an estimate for \( x(t_0 + h/2) \), the \( x \) value at the midpoint in time. We approximate \( x \) at the midpoint using our first order Euler method

\[
x(t_0 + h/2) \sim x(t_0) + h/2f(x_0, t_0) + O(h^2)
\]

We want to insert this into equation 3. Because

\[
x(t_0 + h/2) \sim x(t_0) + h/2f(x_0, t_0)
\]

are equal to first order in \( h \)

\[
hf(x(t_0 + h/2), t_0 + h/2) \sim hf(x(t_0) + h/2f(x_0, t_0), t_0 + h/2)
\]

are equal to second order in \( h \). This means that when we can sub in \( x(t_0) + h/2f(x_0, t_0) \) for \( x(t_0 + h/2) \) in equation 3 the second order terms are equivalent and so the approximation is still good to second order. And consequently the error still \( O(h^3) \).

This approximation using the midpoint is also called a second order Runge-Kutta method. We can also write the method as follows

\[
\begin{align*}
k_1 &= hf(x_0, t_0) \\
k_2 &= hf(x_0 + k_1/2, t_0 + h/2) \\
x(t_0 + h) &= x(t_0) + k_2
\end{align*}
\]
Figure 2: Using the midpoint the estimate for the slope can be improved. A second order Runge-Kutta estimate.
1.2.1 The 4th order Runge-Kutta method

The idea of using midpoints to improve the order the computation it is possible to make a higher order approximation, still using only the ability to evaluate the function for different \(x, t\) values. The Fourth order Runge-Kutta looks like this

\[
\begin{align*}
    k_1 &= hf(x_0, t_0) \\
    k_2 &= hf(x_0 + k_1/2, t_0 + h/2) \\
    k_3 &= hf(x_0 + k_2/2, t_0 + h/2) \\
    k_4 &= hf(x_0 + k_3, t_0 + h) \\
    x(t_0 + h) &= x(t_0) + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4)
\end{align*}
\]

Using Taylor expansions it is possible to show that this is accurate to 4th order so the error \(\propto h^5\). For a few more computation steps, the fourth order Runge-Kutta vastly improves the accurate of the integration. Due to its simplicity and accuracy the 4th order Runge-Kutta integrator is a workhorse of solvers.

Suppose we desire the most accurate possible integration. The order of the integrator \(n\) relates the size of the error to the step size, \(e_h \propto h^{n+1}\). We set the error in each step to be the minimum possible given the double floating point precision level of \(\sim 10^{-16}\) and this lets choose our step size \(h\), with

\[h \sim e_h^{\frac{1}{n+1}}\]

The number of steps we need to integrate an interval \(T\) is

\[N = T/h \propto e_h^{\frac{1}{n+1}}\]

For a second order integrator \(N \propto e_h^{-1/3}\) and for a fourth order integrator \(N \propto e_h^{-1/5}\). The minus third power of \(10^{-15}\) is \(10^5\) but the minus fifth power of \(10^{-15}\) is 1000. Even if we need 10 times the number of computations for each step in the fourth order integrator, it will likely take less time to do the total integration than if we use a second order method.

A large number of small steps are needed to integrate a function that is rapidly changing. If the function is not continuous (or stiff) then integrator routines can give an error. An example of a stiff setting might be a mass on a frictional surface where the force on the body is not a smooth function of velocity. In such a setting you might approximate your force with a continuous form so that your integrator is better behaved.

1.3 Adaptive Step-Sizes

Looking at the error in the first order Newtonian method for integration of an ordinary differential equation

\[\text{Err} \sim f'(x_0)h^2/2\]
If the slope is large then the error is large. We might want to adjust our step size so that we achieve the same level of accuracy at all times.

An example setting might be an eccentricity Keplerian orbit where the body is moving fast nearing pericenter and slowly at apocenter. To integrate the orbit carefully you would need to take small stepsizes (in time) at pericenter but at apocenter, where the body is moving slowly, one could take extremely large stepsizes without compromising in accuracy.

To implement this we take two steps of $h$ and compare the result to what we get if we take a single step of $2h$. Suppose we get $x_1$ after taking the single large step and $x_2$ after taking the two smaller steps. The difference between the two computations tells us the size of the error. If the error is very small we can increase the step size before we take the next step. And if the error is too big we would reduce the step size and try again, only taking the step size if we are happy with the size of the error.

A single step of $2h$ gives an error $\propto (2h)^{n+1}$ where $n$ is the order of the integrator. Whereas a single step of $h$ gives an error of size $\propto h^{n+1}$. It would be nice to add two of these in quadrature but errors could add during consecutive timesteps. We compare

$$(2h)^{n+1} \quad \text{vs} \quad 2h^{n+1}$$

The error is much larger for the single large step than for the two smaller steps. This means that $|x_2 - x_1|$ essentially measures the size of the error for the single large timestep of size $2h$.

There is an unknown constant $c$ relating the size of an error to the step size. For a single step of size $h$

$$\text{err} = ch^{n+1}$$
We can estimate the coefficient $c$ using the difference $|x_2 - x_1|$

$$|x_2 - x_1| \sim c(2h)^{n+1} \tag{4}$$

Solving for $c$

$$c \approx \frac{|x_2 - x_1|}{(2h)^{n+1}}$$

If our desired level of accuracy per unit time is $\delta$ then we want to choose our new stepsize $h'$ so that

$$\text{err} = ch'^{n+1} \sim \delta h'$$

or

$$h' \sim \left(\frac{\delta}{c}\right)^{\frac{1}{n}}$$

Taking our estimated value for $c$ and put it in this equation we find

$$h' = \left(\frac{\delta}{|x_2 - x_1|}\right)^{\frac{1}{n}} (2h)^{1+\frac{1}{n}}$$

And this gives us a way to adjust the step size to the desired level of accuracy.

An adaptive step size integrator is more computationally expensive than a fixed step size integrator. However, you would speed up the calculation overall by taking larger steps when the system is slowly varying. Furthermore if you by mistake take steps that are too large when the solution is rapidly varying you might get an unstable or unbelievable solution. For single particle dynamics one preferentially chooses a variable step size integrator.

1.4 More than one degree of freedom and dynamics with a potential force

Much of the above discussion can be generalized for multidimensional systems where $x$ is a vector.

For example Newtonian dynamics in one dimension, the acceleration $\ddot{y}$ is equal to a force divided by a mass

$$m\ddot{y} = f(y)$$

and we can make this look like a dynamical system with

$$x = \begin{pmatrix} y \\ \dot{y} \end{pmatrix}$$

and

$$\dot{x} = F(x)$$
with
\[
\mathbf{F}(\mathbf{x}) = \left( \begin{array}{c} \dot{y} \\ f(y)/m \end{array} \right)
\]

If we have a potential force then the force
\[
f(y) = -\frac{\partial U(y)}{\partial y}
\]
where \(U(y)\) is the potential energy.

It is convenient to work with potential energy per unit mass and force per unit mass, so henceforth we set \(m = 1\). Let \(\dot{y} = v\). We compute
\[
\begin{align*}
\dot{v} &= -\frac{dU(y)}{dy} \\
v\ddot{v} &= -\frac{dU}{dy}v = -\frac{dU}{dy}\dot{y} = -\frac{d}{dt}U(y) \\
\frac{d}{dt}\frac{v^2}{2} &= -\frac{d}{dt}U(y) \\
\frac{v^2}{2} + U(y) &= \text{constant}
\end{align*}
\]
We recognize energy as a conserved quantity.

However when we call our integrator, it is not obvious whether energy is conserved to high order. In fact, it is unlikely. The integrators we have discussed are designed to match the equations of motion to a certain order and so to low order in step-size they will conserve energy. Using a first order Newton method or a Runge-Kutta method, we have not required energy to be conserved at each step, so on long timescales there will be a drift in energy.

1.5 The meaning of incompressible flow

Let us expand on the connection between being divergence free and being incompressible. Consider a distribution of particles in phase space. The number of particles per unit volume would be \(\rho(p, q)\) or \(\rho(x)\). This the number of particles at position \(x\) in a region with volume \(dV = dx^{2N} = dp^N dq^N\). We can consider the flux \(\rho \mathbf{v}\) of particles out of each surface of a box that has volume \(dV\). Integrating the flux over all surfaces of a box gives us the number of particles leaving the box per unit time.
\[
\frac{dN}{dt} = \int_S \rho \mathbf{v} dA
\]
over the surface, $S$, of the box. Here the velocity $\mathbf{v} = \dot{\mathbf{x}}$. Using Gauss’ law we can rewrite this as

$$\frac{dN}{dt} = \int_{V} \nabla \cdot (\rho \mathbf{v}) dV$$

If particles are leaving the box then the number density in the box must decrease.

$$\frac{dN}{dt} = - \int_{\partial V} \frac{\partial \rho}{\partial t} dV$$

We find a conservation law

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0$$

which can be recognized as conservation of mass for a fluid flow. Consider a bowling ball moving through water. As the bowling ball moves past an observer, it would appear that the density at some location is changing. However the bowling ball is not being compressed. The problem is that the above equation is written respect to a fixed coordinate system and we need to think about the density in a volume element that is moving with the fluid.

We can rewrite this equation as

$$\frac{\partial \rho}{\partial t} + (\mathbf{v} \cdot \nabla) \rho = -\rho \nabla \cdot \mathbf{v}$$

$$\frac{D \rho}{Dt} = -\rho \nabla \cdot \mathbf{v}$$

(5)

where I have used

$$\frac{D}{Dt} \equiv \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla$$

is the advective derivative or that moving with the fluid. Consider $\rho(\mathbf{x}, t)$ as it varies

$$\frac{D \rho}{Dt} = \frac{\partial \rho}{\partial x} \dot{x} + \frac{\partial \rho}{\partial y} \dot{y} + \frac{\partial \rho}{\partial z} \dot{z} + \frac{\partial \rho}{\partial t}$$

$$= \mathbf{v} \cdot \nabla \rho + \frac{\partial \rho}{\partial t}$$

The equation 5 implies that when $\nabla \cdot \mathbf{v} = 0$ then $\frac{D \rho}{Dt} = 0$ and the density of a distribution of particles remains fixed even as the distribution of particles is deformed.

Let us think about this in another way. What does it mean to have $\nabla \cdot \mathbf{v} = 0$ in a uniform density medium? In one dimension this means sense as particles never bunch together and the gradient of $v$ is zero.
Figure 4: On left we show trajectories in a system where area is preserved. In the middle we show a system where volume contracts, as would occur when energy dissipation takes place. On the right the system gains energy.

1.6 Volume conservation in Phase Space

In a two-dimensional system on the plane with density $\rho(x, y, t)$ conservation of mass is

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot (\rho \mathbf{v})$$

with $\mathbf{v} = (\dot{x}, \dot{y})$. We can show this by thinking of mass flux. We can also use Gaus’s law to show this.

If the system is incompressible and density stays the same at all times then

$$\nabla \cdot \mathbf{v} = 0$$

or

$$\partial_x v_x + \partial_y v_y = \partial_x \dot{x} + \partial_y \dot{y} = 0$$

Now going back to our 2d integrated system with $\mathbf{x} = (y, v)$. The condition of incompressibility is

$$\partial_y \dot{y} + \partial_v \dot{v} = 0$$

Recall that $\dot{y} = v$ and so $\partial_y \dot{y} = 0$. Also $\dot{v} = -\partial_y U(y)$ is independent of $v$ and so $\partial_v \dot{v} = 0$.

We can think of our two dimensional system as conserving density in phase space or being incompressible in phase space. This is related to energy conservation and that we have used a conservative force. If the force depends on velocity, then $\partial_v \dot{v} \neq 0$ and volume in phase space is not conserved. In a dissipating system, volume contracts and in a system that gains energy volume increases. See Figure 4.
1.7 Comparison between symplectic and non-symplectic first order integration for the Harmonic oscillator

A simple example is that of a harmonic oscillator. Setting the spring constant $k = 1$ and momentum per unit mass $p = \dot{q}$, the energy per unit mass is

$$H = \frac{1}{2}(p^2 + q^2)$$

The force per unit mass is $f(q) = -q$ and potential energy $U(q) = \frac{q^2}{2}$. The equations of motion are

$$\dot{q} = p$$
$$\dot{p} = -q$$

and together $\ddot{q} = -q$ giving us oscillation with angular frequency $\omega = 1$.

The exact evolution of the system is

$$\begin{pmatrix} q(\tau) \\ p(\tau) \end{pmatrix} = \begin{pmatrix} \cos \tau & \sin \tau \\ -\sin \tau & \cos \tau \end{pmatrix} \begin{pmatrix} q(0) \\ p(0) \end{pmatrix}$$

We note that

$$\left| \det \begin{pmatrix} \cos \tau & \sin \tau \\ -\sin \tau & \cos \tau \end{pmatrix} \right| = \cos^2 \tau + \sin^2 \tau = 1$$

The matrix is the Jacobian of the transformation. Why is this? Think of the transformation as $x' = Ax$. The Jacobian is a matrix composed of elements $\frac{\partial x'}{\partial x_j}$ and as the relation between $x'$ and $x$ is linear $\frac{\partial x'}{\partial x_j} = A_{ij}$. So the matrix is the Jacobian matrix of the transformation. And consequently if $| \det A | = 1$ the transformation is volume preserving. Another way to think about volume preservation is

$$\int_V dx^n = \int_{V'} | \det J | dx'^n$$

If we expand the cosine and sine to first order in $\tau$ we can find a first order approximation to the evolution

$$\begin{pmatrix} q' \\ p' \end{pmatrix} = \begin{pmatrix} 1 & \tau \\ -\tau & 1 \end{pmatrix} \begin{pmatrix} q \\ p \end{pmatrix}$$

However

$$\det \begin{pmatrix} 1 & \tau \\ -\tau & 1 \end{pmatrix} = 1 + \tau^2 \neq 1$$

After one timestep the energy is

$$H' = \frac{1}{2}(p'^2 + q'^2) = \frac{1}{2}(1 + \tau^2)(p^2 + q^2)$$
Since $1 + \tau^2$ must be greater than zero, the energy will increase every timestep. Volume in phase space has not been conserved. After many timesteps the trajectory will spiral outwards. Note that the determinant of the matrix in equation (1.7) is equal to $1 + \tau^2$ and is greater than 1 implying that the area increases.

A symplectic or area preserving in phase space scheme can be constructed with

$$\left( \begin{array}{c} q' \\ p' \end{array} \right) = \left( \begin{array}{cc} 1 & \tau \\ -\tau & 1 - \tau^2 \end{array} \right) \left( \begin{array}{c} q \\ p \end{array} \right)$$

(6)

The determinant of the transformation matrix is 1. The transformation is area preserving and so symplectic. Since the transformation is symplectic it does preserve an energy. However the quantity that is conserved is not the original energy $H$. The transformation gives

$$q' = q + \tau p$$
$$p' = -\tau q + (1 - \tau^2)p$$

giving energy after a single timestep of

$$2H' = q^2 + p^2 + \tau^2(q^2 - p^2) + \tau^4p^2 + 2\tau^3qp$$

Energy is not conserved. However by inserting $q', p'$ into this

$$H_{\text{integrated}} = \frac{1}{2}(p^2 + q^2) + \frac{\tau}{2}pq$$

(7)

we can show that this quantity is conserved and does not change. The integrated Hamiltonian preserved differs from the true Hamiltonian. The above Hamiltonian is called the “modified Hamiltonian.” Just as finite differencing techniques can better approximate a modified differential equation, a symplectic method preserves a Hamiltonian that differs from that intended.

$$H_{\text{integrated}} = H_{\text{true}} + H_{\text{error}}$$

where $H_{\text{true}}$ is that for the harmonic oscillator, $H_{\text{integrated}}$ is that in equation (7) and the difference depends on the timestep

$$H_{\text{error}} = \frac{\tau}{2}pq$$

Since the $H_{\text{integrated}}$ is conserved, it is likely that the difference $H_{\text{integrated}} - H_{\text{true}}$ never gets very big. In this sense we expect the error to be bounded and not grow forever.

The integrated Hamiltonian (equation 7 depends on the timestep used. If the timestep is changed then the integrated Hamiltonian changes. That means adaptive step-size integrators do not preserve a Hamiltonian (approximate to the true one) and so are no longer symplectic. This presents a limitation for symplectic integrators.
1.8 Comparison between 2-nd order Runge Kutta and leap-frog midpoint methods

The second order Runge-Kutta estimates the slope at the step midpoint and then advances to the next time. It always uses the position at beginning of each timestep to compute the mid-point slope. In contrast the leap-frog method uses the previous midpoint position to estimate the position of the next midpoint. The slopes are taken from the midpoint to compute the full step and from the full step to advance the midpoints.

Figure 5: Illustrating the difference between midpoint methods. Both methods are second order, however the leapfrog is time reversible and the second order Runge-Kutta method is not.

Here is our previous midpoint (2nd order Runge-Kutta method)

\[
\begin{align*}
    x(t + h/2) &= x(t) + \frac{1}{2} hf(x, t) \\
    x(t + h) &= x(t) + hf(x(t + h/2, t + h/2))
\end{align*}
\]

The leap-frog method does this

\[
\begin{align*}
    x(t + h) &= x(t) + hf(x(t + h/2, t + h/2, t + h/2)) \\
    x(t + \frac{3}{2} h) &= x(t + h/2) + hf(x(t + h, t + h))
\end{align*}
\]

Both methods are second order. However the leap-frog method has an advantage, it is time symmetric, which means that you integrate forwards and then backwards and get back to where you started. Systems that conserve energy or volume in phase space are time symmetric. Whereas dissipative systems are not.
1.9 Integrating many particles

Consider the setting where there are a number of bodies interacting, for example massive bodies interacting via gravity (this is called N-body). To carry out an integration we need to compute accelerations on every particle. That means we need to know where every particle is when computing accelerations. One way to do this is to integrate all particles with the same step-size. It can be challenging to use a variable step size integrator, so more commonly a low order integrator is used on all particles during each timestep. The step size can be chosen so that it is appropriate for the particle with the largest acceleration. A single particle getting close to another particle can slow down an entire integration. If the step size is too large, then when a particle gets close to another particle it will see all of a sudden a huge force and it can be ejected from the system. This is particularly a problem if the force law is steep (such as for the Lennard-Jones potential).

When doing an N-body integration how would we know if we have an accurate simulation? We can check energy conservation summing all particles. We can check that total angular momentum is roughly conserved. When running the integration there will be drifts in both quantities and we will have to decide if they are preventing us from understanding the modeled system. Inaccuracy in the integration will give noise that should not be there. N-body integrations are extremely sensitive to initial conditions (have short Lyapunov exponents). When the number of particles is more than a few, N-body simulations are swiftly diverging systems, so rapidly lose memory of their initial conditions. An error is made during each timestep but we can ask: Is the system giving a trajectory that exists if we could integrate it exactly? If so then we don’t necessarily need to worry about the inaccuracy of the integration. Arguments like these were hotly debated during the 80s when people started doing N-body integrations and wanted to interpret their results as being relevant for systems like galaxies.

1.10 Burlisch-Stoer Method

Extrapolation!

1.11 Low Order Symplectic integrators – Störmer-Verlet (leapfrog)

For a simple evolution equation
\[
\frac{dz}{dt} = f(z)
\]

An Euler method is
\[
z^{n+1} = z^n + \tau f'(z^n)
\]

It is explicit rather than implicit. For a Hamiltonian system we need to update both momenta and positions. The order of the update is important. For first order integrations if the coordinates are updated first the method is called the assymetric Euler A method. If the momenta are updated first the method is called the Euler B method.
For a separable Hamiltonian

\[ H(p, q) = p^T M p + V(q) \]

with symmetric inverse mass matrix \( M \). Hamilton’s equations give

\[ \dot{q} = M p \quad \dot{p} = -\nabla_q V(q) \]

The Störmer-Verlet method

\[
\begin{align*}
p^{n+\frac{1}{2}} &= p^n - \frac{\tau}{2} \nabla_q V(q^n) \\
q^{n+1} &= q^n + \tau M p^{n+1/2} \\
p^{n+1} &= p^{n+\frac{1}{2}} - \frac{\tau}{2} \nabla_q V(q^{n+1})
\end{align*}
\]

Momenta are evaluated at half steps. When the momenta are eliminated altogether the method can be called the leapfrog method. It turns out this method is 2nd order and symplectic. The above is a series of transformations each which is symplectic. The first and last transformations evolve the potential energy Hamiltonian by \( \tau/2 \). The middle transformation evolve the kinetic energy term by \( \tau \). Each piece is a symplectic transformation so the series is a symplectic transformation. We will show below that if the Hamiltonian is partitioned into any two pieces that the above procedure is a second order approximation.